Effects of Regional Dialect on Word-Final Consonant Voicing

Adam J. Royer

April 9, 2014

Abstract

Two cues that signal phonological voicing in word-final obstruents in English are the amount of glottal pulsing during the consonant and the ratio of the duration of the vowel to the duration of the consonant. Purnell et al. (2012) examined the realization of the voicing contrast in word-final obstruents in two varieties of Wisconsin English and found that both of these cues differed across the two varieties. Smiljanic and Bradlow (2008) found that the temporal cue did not vary between clear speech and plain lab speech, but did not examine spontaneous speech. This study builds on these findings and examines the roles that temporal and non-temporal cues play in signalling the phonological voicing feature of word-final obstruents as a function of regional dialect (Experiment 1), speaking style (read speech vs. spontaneous speech) (Experiment 2), and manner (stop vs. fricative) (Experiment 3). The results suggest that the phonological contrast is maintained phonetically by both cues across dialects, speaking styles, and manners. However, spontaneous speech appears to produce a reduction in the contrast for the temporal cue compared to the contrast maintained in read speech. In addition, the Southern talkers make a smaller distinction between the phonological types than Northern and Western talkers. Therefore it seems that there is regional dialect variation across regional dialects and that more research is warranted into how this variation manifests across voicing cues.

Introduction

Studies of regional American English dialects focus mostly on variation in vowels (Labov et al., 2006; Thomas, 2001), leaving questions unaddressed as to the type and extent of consonantal variation across regional dialects. However, in one study examining consonant variation, Byrd (1994) examined the frequency of word-final stops becoming alveolar flaps, among other measures that weren't strictly about consonants. Byrd found that dialect did affect the frequency of flapping. Northern and North Eastern talkers flapped less than the other regional American English dialects and Northern Midland talkers flapped the most. Byrd's work demonstrates that not only are vowels an important source of variation across regional dialects, but consonants merit research into their variation as well.

One of the most commonly examined phenomena in consonantal variation is obstruent voicing. In English, it has been studied from a variety of angles: word-initial and word-medial stops (Flege and Brown Jr, 1982), different Englishes from around the world (Docherty, 1992; Holmes, 1996), and ethnic varieties of American English (Farrington, 2012).

Purnell et al. (2005, 2012) investigated the extent to which the realization of voicing of stops and fricatives varied in two varieties of Wisconsin English. The first of these varieties was located in a region historically settled by German immigrants and the other was heavily Anglo-American. Purnell et al. hypothesized that because German is a voicing neutralizing language. A neutralizing language is one that minimizes the phonetic distinctions between the phonological categories (i.e. p -> p / _# ; b -> p / _#). This characteristic could have carried over from German speakers to the English of Wisconsinites living in that substantially German-settled area. It was found that there were differences in the signalling of voicing between the two varieties of Wisconsin English for both vowel duration to consonant duration ratio (V/C duration ratio) as well as glottal pulsing during the consonant. The Wisconsinites from the Anglo-American area exhibited larger phonetic distinctions between phonologically voiced and voiceless stops and fricatives while the other group had much smaller distinctions. Purnell et al.'s findings point to potential regional variation, in the extent to which these two cues are used to signal the voicing distinction.

Flege and Brown Jr (1982) looked at the duration of stop closures in bi-syllabic words like "baba" and "papa" in an attempt to see the effect of word/syllable position on the phonetic realization of voicing. Flege and Brown found that word-medially, voiceless stop closures were longer than voiced stop closures. Word-initially this duration difference disappeared as the voiced stop closures became longer to match the durations of the closures of the voiceless stops.

With regard to ethnic varieties of English, Thompson (1975a) studied Mexican-American English by examining word-final devoicing of the phoneme /z/ to [s]. It should be noted that Thompson's categorization of [s] over [z] was based on perception, as he made no mention of any phonetic analysis of the sounds to verify glottal pulsing. He found that a variety of sociolinguistic factors, such as socioeconomic status and education, negatively correlated with the usage of [s] for /z/. That is, talkers who had not reached a high school level of education and talkers of a lower social class tended to devoice /z/ more often than those with higher education and who were of the upper class. Those talkers tended to use [z] more often.

One additional aspect to take note of when examining the voicing characteristic of word-final obstruents in English is the effect that speaking style has on the way that the contrast is maintained. Smiljanic and Bradlow (2008) examined the voicing feature distinction of obstruents in clear speech (i.e. hyperarticulated speech) and plain lab speech. They found that the style of the speech did not affect the duration of the preceding vowel preceding the voiced or voiceless stop in sentence-final position. An interaction between phonological voicing category and style revealed, however, a larger disinction between the voiced and voiceless stops in the plain lab speech than in the clear speech. In their examination of style, however, they did not include any casual or spontaneous speech as a comparison. This comparison is of interest because phonetic reduction of the speech signal is known to happen in casual speech (Dalby, 1986). It seems reasonable to predict that we would see a smaller distinction between the voicing categories for spontaneous speech than for read speech because of this kind of casual speech reduction process.

These studies tell us that the realization of phonological voicing can vary depending on things like the word or sentence position of a segment, the amount of contact a talker has had with other languages, or the dialect of English spoken.

Phonetic Cues to Phonological Voicing

Two types of phonetic cues are used to signal the voicing feature, temporal and non-temporal cues. For the purpose of this study, a temporal cue involves the timing or duration of a phonetic unit, such as the duration of a vowel or frication. The term non-temporal refers to parts of the signal that do not necessarily have to do with a durational measure, such as the presence or absence of glottal pulsing during a consonant (see Figure 1).

Examinations of temporal cues to voicing have found that phonologically voiced obstruents are shorter in duration than voiceless obstruents and also that the vowel immediately preceding a voiced obstruent is longer in duration than the vowel preceding a voiceless obstruent.

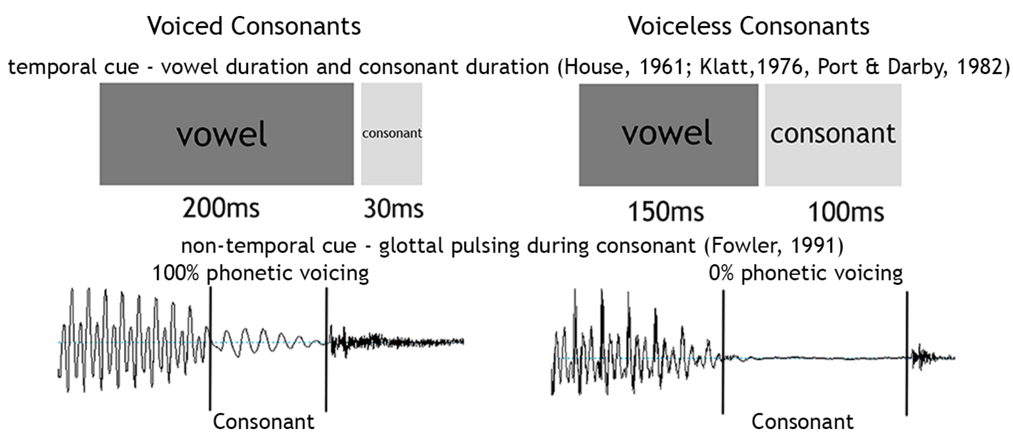

Figure 1: Examples of the temporal and non-temporal cues.

Klatt (1976) argued for a number of distinctions between segments that the temporal cue was the primary perceptual cue (e.g. voiced versus voiceless fricatives). The results he used to support his claim came from Thompson (1975b) who found that shortening the duration of a voiceless fricative could lead a participant to report hearing a voiced fricative.

Raphael (1972) showed that when vowel duration in synthetic CVC(C) words was varied, participants changed their responses as to whether a consonant (or cluster) was voiced or voiceless. The fricatives, stops, and clusters following vowels with a long duration were perceived as being voiced and those same consonants following shorter duration vowels were perceived as being voiceless. The only exception was for the /ʃ/-/ʒ/ distinction where participants' responses varied widely, most likely since English has few instances of minimal pairs of /ʃ/-/ʒ/ in final position.

Another study that points to the importance of the temporal cue in perception of the voicing distinction was done by Port and Dalby (1982). Port and Dalby had participants listen to bi-syllabic non-words and mark on a sheet of paper which of two words they heard (e.g. *dipper* vs. *dibber*). The results of this study suggested that when the silent medial consonant closure was lengthened, participants reported hearing a voiceless consonant more often than they reported hearing a voiced consonant. This increased closure duration made that consonant more prototypically voiceless as Flege's work suggested.

Lisker (1986), however, argued that no singular cue by itself signals voicing but rather that there are multiple cues that work in tandem. Lisker examined the voicing feature for stops in word medial position (i.e. *rabid* vs. *rapid*) and identified 16 different acoustic cues (see Figure 2) as possibly playing roles in a listener's perception of voiced versus voiceless word-medial stops. Lisker cites two of his previous papers (Lisker, 1957, 1981) to support his claim that listeners do not rely on a single cue alone to determine whether a stop is voiced or voiceless. These studies showed that a phonologically voiced stop does not necessarily need a phonetically voiced closure to be reported as phonologically voiced, and a phonologically voiceless stop does not necessarily need to have a long closure duration to be reported as phonologically voiceless.

Thus, while the temporal cue is an important cue in the perception of voiced versus

- Pre-closure
  1. duration of vowel
  2. duration of the first-formant ($F1$) transition
  3. $F1$ offset frequency
  4. $F1$ transition offset time
  5. timing of voice offset
  6. fundamental frequency (F_0 contour)
  7. decay time of signal
- Closure
  8. duration of closure
  9. duration of glottal signal
  10. intensity of glottal signal
- Post-closure
  11. release burst intensity
  12. timing of voice onset
  13. onset of $F1$ transition
  14. $F1$ onset frequency
  15. $F1$ transition duration
  16. F_0 contour

Figure 2: Lisker's (1986) 16 acoustic cues to word-medial stop voicing

voiceless consonants, there are still many other cues that are at work when it comes to the distinction between phonological categories. It is therefore not only important to examine a temporal cue like the vowel to consonant duration ratio, but also other cues like the non-temporal cue of glottal pulsing during a consonant in order to see the complete picture of the phonetic realization of phonological voicing.

The goal of this study was to expand on the previous work on the use of the temporal and non-temporal cues to phonological voicing across regional American English dialects and speaking style. This study extends the dialects of interested from Purnell et al.'s Wisconsin talkers to broader regions. It also includes two speech styles (i.e. read vs spontaneous speech)

Experiment 1 examined the extent to which the primary regional dialects of American English vary in their usage of the temporal and non-temporal cues to voicing for word-final stops. Experiment 2 examined the extent to which Northern, Southern, and Western talkers vary in their usage of the temporal and non-temporal cues to voicing for word-final stops across read and sponanteous speech. Experiment 3 examined the extent to which Northern, Southern, and Western talkers vary in their usage of the temporal and non-temporal cues to voicing for word-final stops and fricatives.

1 Experiment 1

1.1 Methods

1.1.1 Talkers

The data come from recordings of 60 talkers (ages ranging from 18 to 24 years old) recruited from the Indiana University area for The Nationwide Speech Project corpus (Clopper and Pisoni, 2006). There are five male and five female talkers from each of 6 dialect regions were examined: the Mid-Atlantic, Midland, New England, Northern, Southern, and Western. Figure 3 shows the hometowns of each of the 60 talkers.

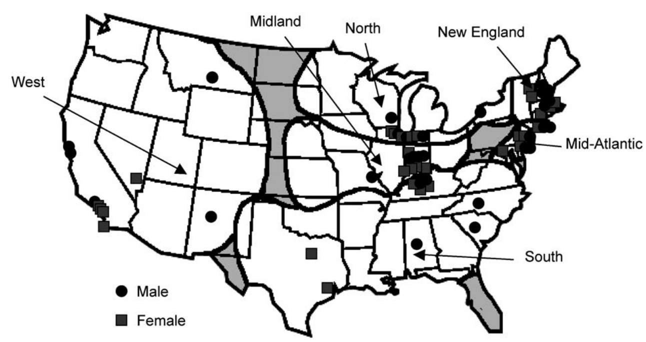

Figure 3: Map of Talkers' Hometowns from the Nationwide Speech Project corpus (Clopper & Pisoni 2006)

1.1.2 Materials

The recordings come from a word list that each participant was asked to read aloud. The word list consisted of 75 CVC words. 24 words ending in stops were examined. 12 of the word-final stops were phonologically voiced and 12 were phonologically voiceless, as shown in Table 1. Of the 1440 tokens, 1129 were analysed in this experiment (311 were excluded because either the wrong word was said, the recording was cut off mid-word, or the stop was unreleased thus making it impossible to mark the stop closure ending). Since only 75 words were available (not all of which ended in a stop), it was not possible to completely balance the tokens based on place of articulation of the coda stop, but the tokens were balanced for voicing as shown in Table 2. The selected word list was also partially balanced for vowel across voicing categories but not place of articulation due to word availability.

voiced CVC tokens	**voiceless CVC tokens**
cab	bite
code	boat
dig	caught
fade	cot
feed	dock
good	doubt
guide	gap
head	keep
loud	lit
mob	poke
tube	tape
void	wet

Table 1: Read Speech Tokens Analyzed

Place of Articulation	voiced	voiceless
bilabial	3	3
alveolar	8	7
velar	1	2

Table 2: Counts of Read Speech Tokens by Place of Articulation

1.2 Analysis

The beginnings and ends of the vowel and the stop were found for each token and marked in a Praat textgrid. The beginning of the vowel was marked at the first sharp increase in energy in the spectrogram and periodicity in the waveform. The end of the vowel and beginning of the stop closure was marked by a drop in amplitude in the waveform and the end of the second formant in the spectrogram. The end of the stop closure was marked at the release burst. On a separate tier, the beginning and ending of glottal pulsing for the vowel and the stop were marked at the beginning and ending of periodicity in the waveform and voice bar in the spectrogram.

The temporal cue was calculated by taking the duration of the vowel and dividing it by the duration of the stop closure. This vowel/consonant duration ratio measurement captures the durational changes to the consonant and its preceding vowel while also accounting for speaking rate variation.

The non-temporal cue was calculated by taking the duration of glottal pulsing during the stop closure and dividing it by the duration of the closure, producing the percentage of the stop closure that was phonetically voiced.

1.3 Results & Discussion

The temporal and non-temporal cue data were analyzed using repeated measures ANOVAs with phonological voicing as a within-subject variable and talker dialect as a between-subject variable. As expected, the phonologically voiced stops have a larger V/C duration ratio than the voiceless stops ($F(1,54)=177.35$, $p<.001$)(see Figure 4) and the phonologically voiced stops have a higher percentage of voicing than the voiceless stops ($F(1,54)=1445.29$, $p>.001$)(see Figure 5). The effect of dialect was not significant for either the temporal cue ($F(1,54)=1.53$, $p=.197$) nor the non-temporal cue ($F(1,54)=2.156$, $p=.0726$). The interactions between phonological type of stop and talker dialect were also not significant for the V/C duration ratio ($F(1,54)=1.248$, $p=.3$) nor the percentage of the consonant that is phonetically voiced ($F(1,54)=1.977$, $p=.0969$).

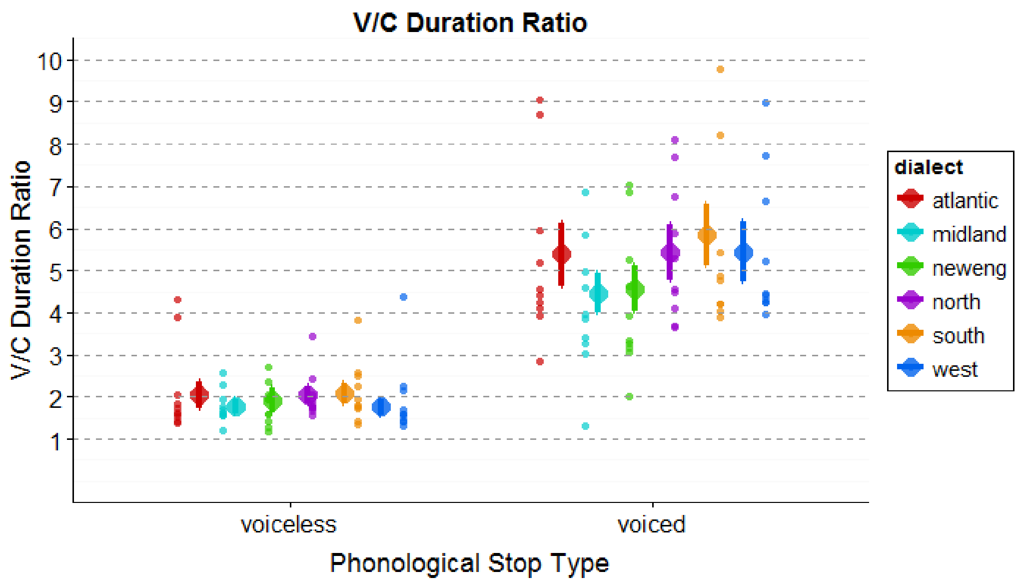

Figure 4: Grand means (large dots) with 95% confidence intervals and subject means (small dots) of the dialects for V/C duration ration for each stop type.

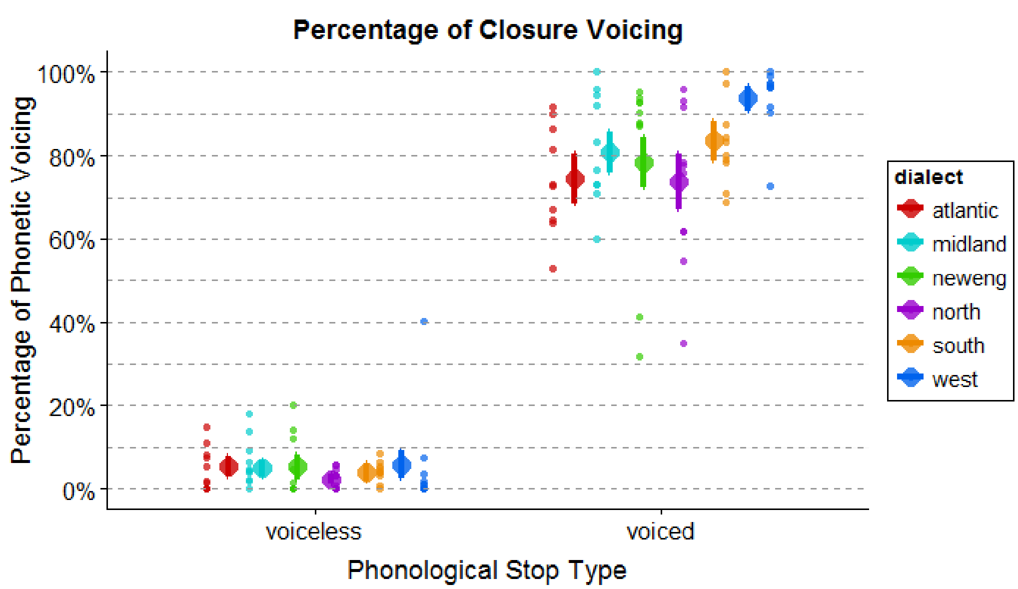

Figure 5: Grand means (large dots) with 95% confidence intervals and subject means (small dots) of the dialects for percentage of closure voicing for each stop type.

These data are consistent with the perceptual data in the previously mentioned studies (Flege and Brown Jr, 1982; Dalby, 1986; Lisker, 1986; Smiljanic and Bradlow, 2008) that show that listeners use the temporal and non-temporal cues to differentiate between voicing categories, so it makes sense that the talkers here would produce the voiced and voiceless stops distinctly. The lack of dialect effect, however, is somewhat surprising and could partly be due to an insufficient amount of data. The average talker produced just 19 usable tokens.

From Figure 5, the dialects appear to vary quite a bit in closure voicing for voiced stops , which is reflected in the dialect effect and interaction with consonant type for this cue approaching significance. It seems prudent, therefore, to narrow the focus of the examinations to just the Northern dialect region along with the most different dialect, the Southern dialect. This comparison could be more useful than a comparison of all 6 dialects at once.

The Northern dialect was chosen as a comparison to Purnell's Wisconsin talkers. The Western dialect was chosen because of its similar temporal properties to the Northern dialect, including speaking rate. The Southern dialect region was chosen because it is too different temporally from the West and North with its longer vowel intervals and slower speaking rate (Clopper and Smiljanic, 2011).

Another dimension on which these cues could vary is speech style, but all of the tokens in this experiment were taken from words read in isolation. Smiljanic and Bradlow (2008) found that style is not a significant factor in vowel duration before sentence-final voiced and voiceless stops but they did find an interaction of phonological type and speech style. In this interaction, the voiceless stops had a lower average vowel duration in the clear speech condition than in the plain lab speech condition, and the voiced stops had a higher average vowel duration in the clear speech condition than in the plain lab speech condition. The two types of speech they looked at, though, were plain lab speech and clear, hyperarticulated speech. Their study did not include spontaneous speech, which might exhibit more reduction than plain lab speech and therefore shed additional light on the interaction which seems to point to an increased distinction between the phonological voicing types. Unlike Smiljanic and Bradlow's study, Experiment 2 includes a spontaneous speech condition to test for the effect of speech style in word-final stop voicing. Because of what we know about phonetic reduction in spontaneous speech, it seems likely that all of the dialects maintain a smaller distinction between voiced and voiceless stops for both cues. We might expect to see a much smaller distinction between voiced and voiceless stops in a spontaneous speech condition than in a read speech condition. In Experiment 2, tokens from the read word lists and interview data available in the corpus were combined to add more data to confirm a dialect effect and allow for the exploration of a speech style effect.

2 Experiment 2

2.1 Methods

2.1.1 Talkers

The same 30 Northern, Southern, and Western talkers from Experiment 1 were used in this experiment.

2.1.2 Materials

The read speech tokens were the same ones used in Experiment 1.

The spontaneous speech tokens were taken from interviews with the 30 talkers. Two types of interviews were recorded. The first type of interview involved the interviewer asking questions to elicit a specific word which also appeared in the question (e.g. "What's the best **date** you've ever been on?"). In the second type of interview, the interviewer asked more open ended and biographic questions about the talker (e.g. "Do you know what you want to do after you graduate?").

The criteria for selecting words for analysis from these interviews were as follows:

1. Tokens were pitch-accented monosyllabic words with a stop coda to ensure that the stop was in the syllable carrying the accent of the target word since there are acoustic differences between consonants in stressed versus unstressed syllables (De Jong, 1995; Fougeron and Keating, 1997). This was also done to maximize similarity to word list tokens.

2. Stops could not be followed by another stop because the closures of the two stops would combine, making it impossible to know where one stop ended and the second one began.

3. Only the first usable instance of a word was used per interview per talker to avoid having too many frequent words for each talker, which would be unbalanced (e.g. 9 instances of "good" versus 1 instance of "shock").

4. Phrase-final stops had to be released so that it was possible to find the end of the stop closure.

The largest number of tokens an individual talker contributed was 14 and the smallest was 2 (mean = 6.8, sd = 3.27). The data were also skewed toward the voiceless tokens (mean = 2.75 voiced tokens per talker, sd = 1.33; mean = 4.05 voiceless tokens per talker, sd = 2.48).

2.2 Analysis

The methods for marking up the stop tokens and preceding vowels and for calculating the temporal and non-temporal cues was the same as in Experiment 1.

2.3 Results & Discussion

The temporal and non-temporal cue data were analyzed using repeated measures ANOVAs with phonological voicing and speech style as within-subject variables and talker dialect as a between-subject variable. Mirroring results from Experiment 1, the phonologically voiced stops have a larger V/C duration ratio than the voiceless stops across dialects ($F(1,27)=97.449$, $p<.001$) (see Figure 6). The interview speech has a significantly smaller V/C duration ratio than the read speech ($F(1,27)=26.810$, $p=.003$). There is also an interaction between phonology and style ($F(1,27)=47.397$, $p<.001$), which is seen in the smaller distinction between the voiced and voiceless stops in interview speech than in read speech. There is no main effect of dialect, however ($F(1,27)=0.08$, $p=.923$).

In an examination of the non-temporal cue data, the phonologically voiced stops have a higher percentage of phonetic voicing than the phonologically voiceless stops ($F(1,27)=609.223$, $p<.001$) (see Figure 7), as expected. The other main effects and none of the interactions are significant.

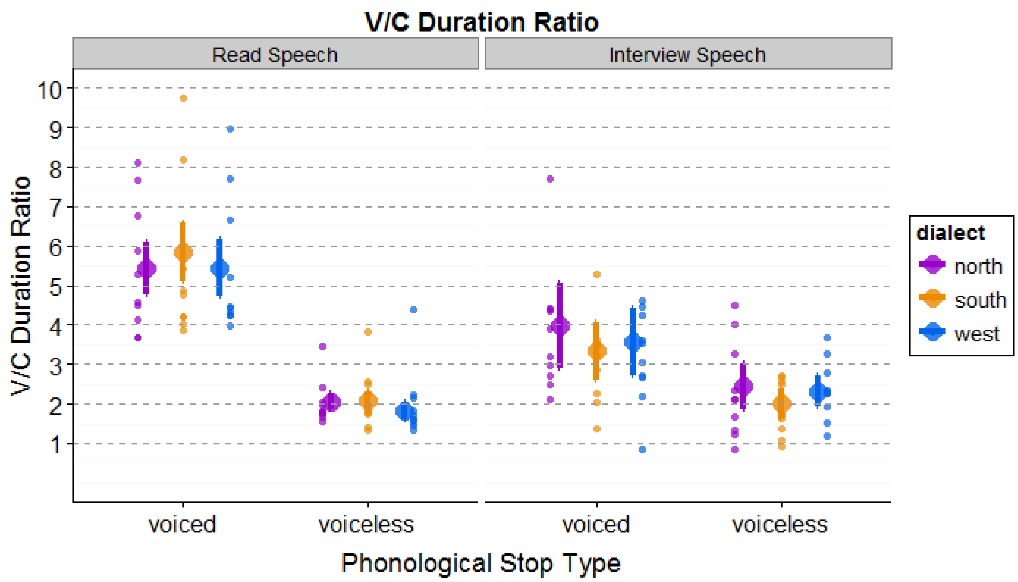

Figure 6: Grand means (large dots) with 95% confidence intervals and subject means (small dots) of the dialects.

The results of Experiment 2 further confirm that both cues are being used to maintain the voicing distinction between the word-final stops across all 3 dialects and both speaking styles. The prediction about the interaction of speech style and phonological voicing category of the stop was observed for the temporal cue only..

One peculiarity does remain in the data, which is that the V/C duration ratio for the Southern talkers is numerically smaller for both voiced and voiceless stops than the Northern and Western talkers' ratios in the interview condition. The Northern and Western dialects' ratios seem to pattern similar to each other in both styles, which is what was predicted, but surprisingly the Southern speakers produced numerically smaller ratios than the other two dialects. This result is surprising because we would

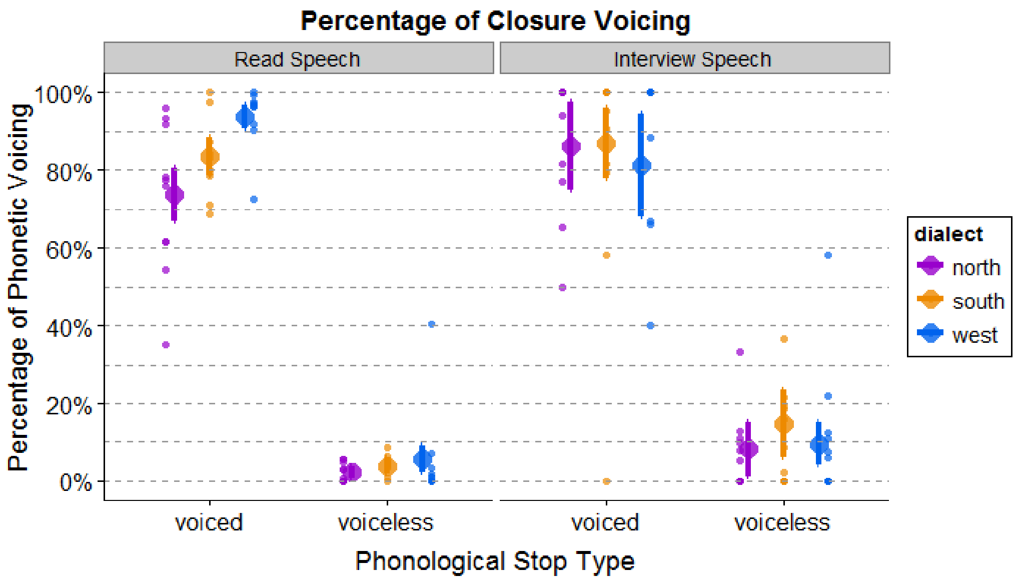

Figure 7: Grand means (large dots) with 95% confidence intervals and subject means (small dots) of the dialects.

expect larger ratios in both styles for Southern talkers since Southern talkers are expected to have longer overall vowel durations than Northern and Western talkers.

When looking at the average contribution of data per talker in the interview condition, it is apparent that the available interview data were sparse. In an effort to better examine the Southern V/C duration ratios and more closely replicate Purnell's examination of stops and fricatives, fricatives from the interview speech are added to the analyses of Experiment 3. The lack of a comparable number of voiced and voiceless fricatives, however, prevents the inclusion of the read speech condition in Experiment 3.

3 Experiment 3

3.1 Methods

3.1.1 Talkers

The same 30 Northern, Southern, and Western talkers from Experiments 1 and 2 were used

3.1.2 Materials

The speech tokens were taken from the same interviews as in Experiment 2. The criteria for selecting tokens were as follows:

1. Tokens were accented monosyllabic words with a fricative or stop coda to ensure that the target segment was in the syllable carrying the accent of the target word since there are acoustic differences between consonants in stressed versus unstressed syllables (De Jong, 1995; Fougeron and Keating, 1997).
2. Consonants could not be followed by another consonant of the same manner because the respective closures of the two stops or frication of the two fricatives would combine, making it impossible to know where one consonant ended an another began.
3. Only the first usable instance of a token was used per interview per talker to avoid having too many frequent words for each talker, which would be unbalanced.
4. Phrase-final stops had to be a released so that it was possible to find the end of the stop closure.

There were total of 551 tokens. The contribution of each talker to the data varied. The largest number of tokens a talker contributed was 29 and the smallest was 8 (mean = 18.37; sd = 6.91). The data were also skewed toward the voiceless tokens (mean = 11.77 voiceless tokens per talker, sd = 5.69; mean = 6.60 voiced tokens per talker, sd

= 3.02). There were also more fricatives than stops (mean = 11.60 fricative tokens per talker, sd = 5.94; mean = 6.77 stop tokens per talker, sd = 3.10).

3.2 Analysis

The method for marking up the stop tokens and calculating the temporal and non-temporal cues was the same as in Experiment 2.

For fricatives, the end of the vowel and beginning of the fricative was marked by a drop in amplitude in the waveform and the appearance of frication in the waveform. The end of the fricative was marked by a sharp increase (e.g. into a vowel) or decrease (e.g. into silence) in amplitude frication in the waveform. In another annotation tier the beginning and ending of glottal pulsing of the fricatives was marked at the beginning and ending of periodicity in the waveform and voice bar in the spectrogram.

3.3 Results & Discussion

The data from the interviews varied and produced uneven contributions of data from particular tokens and from particular talkers. Because of this unbalanced data set, linear mixed effects models were used to test for dialect and obstruent effects instead of the ANOVAs used in Experiments 1 and 2.

Table 3 shows the results of the linear mixed effects models for the temporal cue. Table 4 shows the results of the linear mixed effects models for the non-temporal cue.

The models were first built with maximal design-driven random slopes for both item and talker. These models did not converge, so the smallest models that fit as well as the largest models were chosen. The models presented here had random slopes for dialect for token, random intercepts for token, random slopes for phonological voicing category for talker and random intercepts for talker. All of the random effects were uncorrelated.

The fixed effects included talker dialects, voicing category of consonant, and place of articulation of consonant. Two models were built, one with the Western dialect as the reference level and another model with the Southern dialect as the reference level so as to get the comparisons between all three dialects.

Significance was determined by the absolute value of the model's t-value for a factor or interaction being equal to or greater than 2.

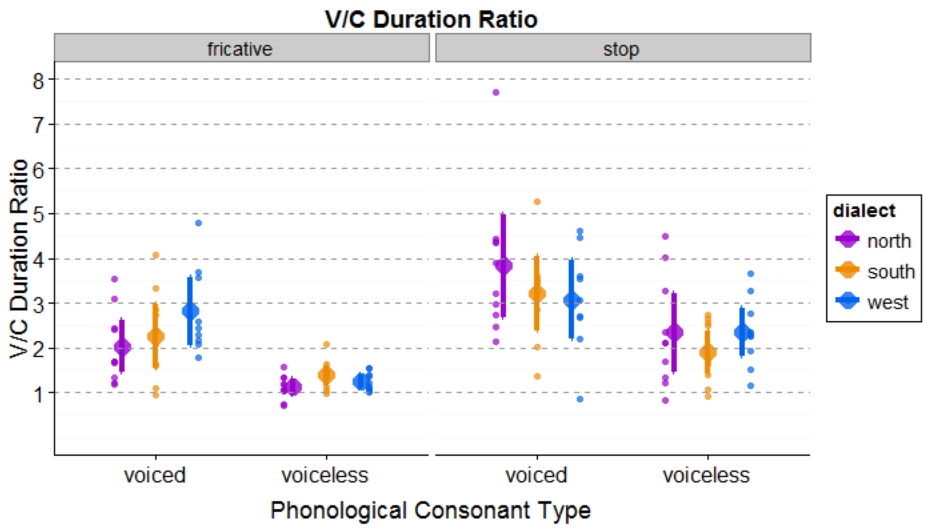

Figure 8: Grand means (large dots) with 95% confidence intervals and subject means (small dots) of the dialects.

Table 3: LMEMs for Vowel Duration to Consonant Duration Ratio
Coefficients std errors in parentheses

	V/C Duration Ratio	
	Reference level - West	Reference level - South
Intercept	2.636 (0.299)	2.272 (0.389)
South	-0.177 (0.357)	
West		0.492 (0.607)
North	-0.380 (0.377)	-0.162 (0.521)
Stop	0.911 (0.473)	**1.270* (0.503)**
Voiceless	**-1.420* (0.361)**	**-0.974* (0.413)**
Stop x Voiceless	0.191 (0.517)	-0.543 (0.515)
South x Stop	0.100 (0.594)	
West x Stop		-0.423 (0.787)
North x Stop	0.878 (0.623)	0.664 (0.636)
South x Voiceless	0.235 (0.417)	
West x Voiceless		-0.562 (0.672)
North x Voiceless	0.392 (0.449)	0.111 (0.544)
South x Stop x Voiceless	-0.408 (0.627)	
West x Stop x Voiceless		0.777 (0.841)
North x Stop x Voiceless	-0.762 (0.678)	-0.283 (0.634)
Observations	551	551
Log Likelihood	-963	-980
AIC	1981	2014
BIC	2097	2130

Significance: *|t|>2

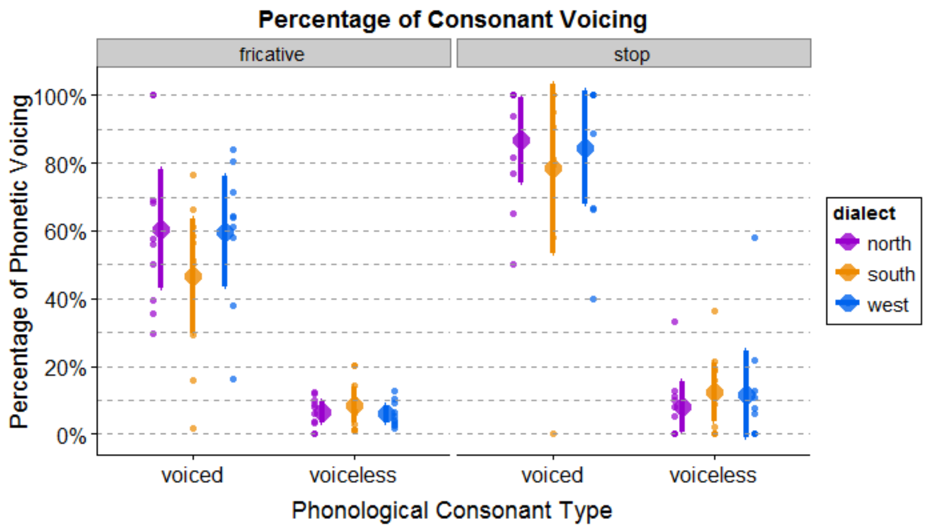

Figure 9: Grand means (large dots) with 95% confidence intervals and subject means (small dots) of the dialects.

Table 4: LMEMs for Percent of Phonetic Voicing During Consonant
Coefficients std errors in parentheses

	Percent Voiced	
	Reference level - West	Reference level - South
Intercept	0.597 (0.058)	0.441 (0.055)
South	-0.157 (0.079)	
West		0.157 (0.079)
North	-0.004 (0.080)	0.153 (0.078)
Stop	**0.227* (0.073)**	**0.406* (0.070)**
Voiceless	**-0.535* (0.065)**	**-0.354* (0.061)**
Stop x Voiceless	**-0.186* (0.078)**	**-0.345* (0.076)**
South x Stop	0.179 (0.100)	
West x Stop		-0.179 (0.100)
North x Stop	0.042 (0.099)	-0.136 (0.096)
South x Voiceless	**0.181* (0.087)**	
West x Voiceless		**-0.181* (0.087)**
North x Voiceless	0.010 (0.089)	-0.171 (0.087)
South x Stop x Voiceless	-0.158 (0.107)	
West x Stop x Voiceless		0.158 (0.107)
North x Stop x Voiceless	-0.062 (0.108)	0.096 (0.106)
Observations	551	551
Log Likelihood	15	15
AIC	12	12
BIC	102	102

Significance: *|t|>2

From the models for the vowel to consonant duration ratio (see Table 3 and Figure 8), it appears that there is a significant difference between the voiced and voiceless obstruents across dialects with the voiced obstruents have higher percentages and V/C duration ratios then the voiceless obstruents. There is also a significant difference between the ratios for fricatives and stops. The main effect of dialect and all of the interactions were not significant.

Modelling the percentage of phonetic voicing (see Table 4 and Figure 9) confirmed that the dialects have significantly different percentages of voicing between voiced and voiceless obstruents as well as between stops and fricatives. The voiced obstruents have more phonetic voicing than the voiceless obstruents, and the fricatives have less phonetic voicing than the stops. There are also two significant interactions, one between manner and phonological type, and one between the dialect and phonological type. The manner and phonological type interaction shows that the distinction between voiced and voiceless fricatives is smaller than the distinction between voiced and voiceless stops. The dialect and phonological type interaction reveals that distinction between voiced and voiceless obstruents is smaller for the Southern dialect than the other two dialects.

The main effect of dialect and the other interactions are not significant.

General Discussion

Both the temporal and non-temporal cues maintain contrasts between phonologically voiced and voiceless obstruents in all dialects in all three of the experiments. These findings seem to support Lisker's claim that not just one cue is responsible alone for signalling the voicing distinction, but rather the convergence of a number of cues. This redundancy of cues is why when one cue is unreliable (e.g. no glottal pulsing during a phonologically voiced stop), listeners still perceive that segment as voiced (Lisker, 1957, 1981) since there are other cues in the signal that do signal the voicing category.

This redundancy of cues, however, does not preclude variation in the extent to which these cues are used. In Experiment 3 we see the Southern talkers' usage of the non-temporal cue patterning differently than the Northern and Western talkers' usage of that cue. Although the Southern talkers did not produce a complete neutralization of the voiced and voiceless obstruents, their distinction between the two categories was significantly different from the Northern and Western talkers'.

It is unsurprising that these temporal and non-temporal cues do not vary drastically from one another in the three experiments because if they did, there would be a lot of confusion between dialects when it comes to the numerous minimal pairs in English that differ in obstruent voicing (e.g. "dog" vs. "dock", "bus" vs. "buzz", etc.). Thus, if perceptually the distinction between the voiced and voiceless obstruents is maintained, what variation there is in the cues to phonological voicing must be rather subtle, which is what is found in this study.

There is also the effect of speech style in Experiment 2, which produced a smaller voicing distinction for the temporal cue in interview speech than in read speech. This reduction process most likely has nothing to do with speaking rate differences across styles because taking the durational ratio accounts for that. Thus, it seems that the talkers aren't trying to maintain the distinction between voiced and voiceless stops in fast, casual speech as much as they do in slower, read speech. This reduction process is creating a compression of the distinction between the voiced and voiceless consonants across dialects.

With regard to Purnell's results, they could not be replicated most likely because his investigation looked at small speech communities within the Northern dialect region, not comparisons of large regional dialects. The small, particular group of speakers in the North that devoice their obstruents is most likely washed out when lumped in with Northern speakers using different cues to a different extent. For this reason it might be worth examining these research questions from a more variationist approach. Taking a look at individual differences like personal background, heritage, and language contact could better inform the types of regional variation that goes on, since the variation seems to be tied heavily to the history of specific smaller regions.

Other future research directions might also include the examination of how the cues to voicing are maintained for fricatives across speech styles as was done for the stops in Experiment 2. Since this corpus was not designed for specifically examining word-final stop voicing, it unfortunately does not have the data needed to support such an examination. A production task with a new set of speakers would have to be done which would provide more control over conditions and likely produce more telling results.

References

Byrd, D. (1994). Relations of sex and dialect to reduction. *Speech Communication*, (15):39–54.

Clopper, C. G. and Pisoni, D. B. (2006). The Nationwide Speech Project: A new corpus of American English dialects. *Speech Communication*, 48:633–644.

Clopper, C. G. and Smiljanic, R. (2011). Regional variation in temporal organization in American English. Columbus, OH. Ilse Lehiste Memorial Symposium: Melody and Meter.

Dalby, J. M. (1986). *Phonetic structure of fast speech in American English*, volume 7. Reproduced by the Indiana University Linguistics Club.

De Jong, K. J. (1995). The supraglottal articulation of prominence in english: Linguistic stress as localized hyperarticulation. *The journal of the acoustical society of America*, 97(1):491–504.

Docherty, G. J. (1992). *The timing of voicing in British English obstruents*, volume 9. Walter de Gruyter.

Farrington, C. (2012). The social distribution of devoicing in urban Southern African American English. In *New Ways of Analyzing Variation 41st Annual Conference*, page Session 16, Indiana University Bloomington.

Flege, J. and Brown Jr, W. (1982). Effects of utterance position on English speech timing. *Phonetica*, 39(6):337–357.

Fougeron, C. and Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *The journal of the acoustical society of America*, 101(6):3728–3740.

Holmes, J. (1996). Losing voice: is final /z/ devoicing a feature of Maori English? *World Englishes*, 15(2):193–205.

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, 59:1208–1221.

Labov, W., Ash, S., and Boberg, C. (2006). *The Atlas of North American English*. Mouton de Gruyter, Berlin.

Lisker, L. (1957). Closure duration and the intervocalic voiced-voiceless distinction in English. *Language*, 33:42–49.

Lisker, L. (1981). On generalizing the *rabid-rapid* distinction based on silent gap duration. *Haskins Laboratories Status Report on Speech Research*, SR-65:251–259.

Lisker, L. (1986). Voicing in English: A catalogue of acoustic features signaling /b/ versus /p/ in trochees. *Language and Speech*, 29:3–11.

Port, R. F. and Dalby, J. (1982). Consonant/vowel ratio as a cue for voicing in English. *Perception and Psychophysic*, 32:141–152.

Purnell, T., Raimy, E., and Salmons, J. (2012). Obstruent change in the Upper Midwest: Wisconsin final laryngeal distinctions. In *New Ways of Analyzing Variation 41st Annual Conference*, Indiana University Bloomington.

Purnell, T., Salmons, J., and Tepeli, D. (2005). German substrate effects in Wisconsin English: Evidence for final fortition. *American Speech*, 80(2):135–164.

Raphael, L. J. (1972). Preceding Vowel Duration as a Cue to the Perception of the Voicing Characteristic of Word-Final Consonants in American English. *The Journal of the Acoustical Society of America*, 51:1296–1303.

Smiljanic, R. and Bradlow, A. R. (2008). Stability of temporal contrasts across speaking styles in English and Croatian. *Journal of Phonetics*, 36(1):91–113.

Thomas, E. (2001). *An Acoustic Analysis of Vowel Variation in New World English*. Number no. 85 in Publication of the American Dialect Society. Duke University Press.

Thompson, R. M. (1975a). Mexican-American English: Social correlates of regional pronunciation. *American Speech*, 50(1/2):18–24.

Thompson, R. M. (1975b). The Perception of Voicing in English Affricates and Fricatives. *Journal of the Acoustic Society of America*, 58:1280–1287.